

1. A computer system for database management comprising, means for storing and updating a first set of data, indexing means for storing and updating a selected index of key values related to the first set of data, the indexing means comprising means for selectively updating the index by incrementally updating the index by incrementally adding key values to the index or by fully rebuilding the index, heuristic determination means for selecting the incremental update of the index, or the full rebuild of the index, for a given second set of data to be added to the first set of data.

2. The computer system of claim 1 in which the heuristic determination means comprises a function which takes as input index meta-data, comprising characteristics of the first set of data, the index, and the second set of data.

3. The computer system of claim 2 in which the database management system is a relational database management system in which the index is stored as a binary tree and in which the index meta-data comprises estimates of the table size of the first set of data, the table size of the second set of data, and the height of the index, whereby the function is empirically defined to select the incremental update of the index or the full rebuild of the index based on the predicted relative efficiencies of the incremental update of the index and the full rebuild of the index.

4. The computer system of claim 3 in which the function determines, for each potential binary tree height of the index, a threshold percentage of the table size of the second set of data to the table size of the first set of data, whereby the threshold percentage value for each binary tree height determines the selection of the incremental index update or the full rebuild index update.

1 5. The computer system of claim 1 in which the heuristic determination means
2 comprises means for the user to specify the selection of the incremental update of the
3 index or the full rebuild of the index.

1 6. The computer system of claim 2 further comprising a means for storing the
2 index meta-data independent of the means for storing and updating the first set of data.

1 7. The computer system of claim 2 in which the means for storing the index
2 meta-data comprises a recovery history file for the first set of data.

1 8. The computer system of claim 1 in which the selected index of key values is
2 one of a plurality of indexes and in which the selected index is the first index on the first
3 set of data.

1 9. The computer system of claim 1 in which the selected index of key values is
2 one of a plurality of indexes and in which the selected index is selected on the basis of the
3 relative sizes of each of the plurality of indexes on the first set of data.

1 10. The computer system of claim 2 further comprising means for generating
2 values for the function of the heuristic determination means, for a specified test range of
3 values for each of the first set of data, the index, and the second set of data.

1 11. The computer system of claim 2, the computer system having one or more
2 CPUs, one or more disks, a sort heap and a database bufferpool, and in which the index
3 meta-data is defined to reflect a subset of the following characteristics: the percentage of
4 free space in the index, the estimated size of the index after both the incremental and the
5 rebuild updates of the index, the width of the average key value in the index, the size of
6 the sort heap and the database bufferpool in the computer system, the number and speed
7 of the CPUs in the computer system, and the number and speed of the disks in the system.

003320" 332330

1 12. A computer system for relational database management comprising,
2 means for storing and updating a first set of data,
3 indexing means for storing as a binary tree, and updating, a selected index of key
4 values related to the first set of data the indexing means comprising means for selectively
5 updating the index by incrementally updating the index by incrementally adding key
6 values to the index or by fully rebuilding the index,
7 heuristic determination means for selecting the incremental update of the index, or
8 the full rebuild of the index, for a given second set of data to be added to the first set of
9 data, the heuristic determination means comprising a function which takes as input index
10 meta-data,
11 the index meta-data being stored separately from the first set of data and
12 comprising estimates of the table size of the first set of data, the table size of the second
13 set of data, and the height of the index,
14 whereby the function is empirically defined to select the incremental update of the
15 index or the full rebuild of the index based on the predicted relative efficiencies of the
16 incremental update of the index and the full rebuild of the index,
17 the heuristic determination means further comprising means for the user to specify
18 the selection of the incremental update of the index or the full rebuild of the index.

1 13. A computer program product for use with a computer comprising a central
2 processing unit and random access memory, said computer program product comprising a
3 computer usable medium having computer readable code means embodied in said
4 medium for managing a database, said computer program product comprising:
5 computer readable program code means for causing a computer to store and
6 update a first set of data,
7 computer readable program code indexing means for causing a computer to store
8 and update a selected index of key values related to the first set of data, the indexing
9 means comprising means for selectively updating the index by incrementally updating the
10 index by incrementally adding key values to the index or by fully rebuilding the index,

[illegible]

1 18. A method for updating a selected index in a computer system for database
2 management, the computer system comprising means for storing and updating a first set
3 of data related to the index, the method comprising the steps of:

[illegible]

20. A method for updating a selected b-tree index in a computer system for relational database management, the computer system comprising means for storing and updating a first set of data related to the index, the method comprising the steps of:

(b) selectively updating the index by incrementally updating the index by incrementally adding key values to the index or by fully rebuilding the index, as selected.

21. The method of claim 20 in which the function comprises the step of
determining, for each potential binary tree height of the index, a threshold percentage of
the table size of the second set of data to the table size of the first set of data, whereby the

3 receiving data records to load into the table;
4 selecting one of a first operation and second operation, wherein the first operation
5 incrementally updates the index on the table as each received data record is added to the
6 table and the second operation rebuilds the index from the table after all the received data
7 records have been added to the table; and
8 using the selected first operation or second operation to update the index with the
9 received data.

1 28. The method of claim 27, further comprising:
2 determining which of the first operation or second operation is more efficient,
3 wherein the first or second operation determined to be more efficient is the selected
4 operation used for updating the index with the received data.

1 29. The method of claim 28, wherein determining which operation is more
2 efficient is a function of a percentage of the received data records to add to the table and
3 characteristics of the index.

1 30. The method of claim 29, wherein the characteristics of the index used in
2 determining which operation is more efficient comprise a size and complexity of the
3 index.

1 31. The method of claim 30, wherein the index comprises a binary tree
2 structure, and wherein a height of the index tree is indicative of the size and complexity
3 of the index.

1 32. The method of claim 28, wherein determining which operation is more
2 efficient further comprises considering at least one of a following factors: an estimated
3 time required to extract index keys from the table; an estimated time to sort the index
4 keys; and an estimated time to rebuild the index from the sorted keys.

1 33. The method of claim 28, further comprising:
2 maintaining a list of threshold values for different index sizes; and
3 using the number of received data records to add to the table to determine a
4 comparison value, wherein determining whether the first or second operation is more
5 efficient is based on the comparison value and the threshold for the size of the index to be
6 updated.

0055-0022-23

1 38. The system of claim 37, further comprising:
2 means for determining which of the first operation or second operation is more
3 efficient, wherein the first or second operation determined to be more efficient is selected
4 to use for updating the index with the received data.

1 45. The system of claim 44, wherein the index comprises a binary tree and
2 wherein the list of threshold values provides one threshold for each of a plurality of

46. The system of claim 43, wherein the first operation is more efficient if the comparison value is less than the threshold value and wherein the second operation is more efficient if the comparison value is greater than the threshold value.

4 receiving data records to load into the table;
5 selecting one of a first operation and second operation, wherein the first operation
6 incrementally updates the index on the table as each received data record is added to the
7 table and the second operation rebuilds the index from the table after all the received data
8 records have been added to the table; and

1 48. The program of claim 47, wherein the program is further capable of
2 causing the processor to perform:

49. The program of claim 48, wherein determining which operation is more efficient is a function of a percentage of the received data records to add to the table and characteristics of the index.

1 55. The program of claim 54, wherein the index comprises a binary tree and
2 wherein the list of threshold values provides one threshold for each of a plurality of
3 different height index binary trees, wherein the threshold selected for comparison with the
4 comparison value is based on the height of the index to update.

